THE JOURNAL OF BIOLOGICAL CHEMISTRY
Vol. 275, No. 16, Issue of April 21, pp. 12214–12222, 2000

*Printed in U.S.A.*

# Transcriptional Regulation of the Divergent *paa* Catabolic Operons for Phenylacetic Acid Degradation in *Escherichia coli**

(Received for publication, December 12, 1999, and in revised form, January 20, 2000)

**Abel Ferrández‡, José Luis García, and Eduardo Díaz§**

*From the Department of Molecular Microbiology, Centro de Investigaciones Biológicas, Consejo Superior de Investigaciones Científicas, 28006 Madrid, Spain*

**The expression of the divergently transcribed *paaZ* and *paaABCDEFGHIJK* catabolic operons, which are responsible for phenylacetic acid (PA) degradation in *Escherichia coli*, is driven by the *Pz* and *Pa* promoters, respectively. To study the transcriptional regulation of the inducible *paa* catabolic genes, genetic and biochemical approaches were used. Gel retardation assays showing that the PaaX regulator binds specifically to the *Pa* and *Pz* promoters were complemented with *in vivo* experiments that indicated a PaaX-mediated repression effect on the expression of *Pa-lacZ and Pz-lacZ* reporter fusions. The region within the *Pa* and *Pz* promoters that is protected by the PaaX repressor in DNase I footprinting assays contains a conserved 15-base pair imperfect palindromic sequence motif that was shown, through mutational analysis, to be indispensable for PaaX binding and repression. PA-coenzyme A (PA-CoA), but not PA, specifically inhibited binding of PaaX to the target sequences, thus confirming the first intermediate of the pathway as the true inducer and PaaX as the only bacterial regulatory protein described so far that responds to an aryl-CoA compound. Superimposed in the specific PaaX-mediated regulation is transcriptional activation by the cAMP receptor protein and the integration host factor protein. These global regulators may adjust the transcriptional output from *Pa* and *Pz* promoters to the overall growth status of the cell.**

Phenylacetic acid (PA)[1] is a central compound to which pollutants, such as styrene and *trans*-styrylacetic acid, as well as other aromatic compounds, such as 2-phenylethylamine, phenylacetaldehyde, and several phenylalkanoic acids with an even number of carbon atoms, converge through different peripheral catabolic pathways (1, 2). Aerobic PA catabolism in *Pseudomonas putida* U and *Escherichia coli* W has been reported to represent a novel aerobic hybrid pathway whose first step is the activation of PA to phenylacetyl-coenzyme A (PA-CoA) by the action of a PA-CoA ligase (1, 2). In *E. coli* K-12 and *E. coli* W, the *paa* genes responsible for PA catabolism are clustered in the chromosome at min 31.0 according to the *E. coli* K-12 linkage map (2). The 14 *paa* genes are organized in three transcription units as follows: two divergently transcribed operons, *paaZ* and *paaABCDEFGHIJK*, encoding the catabolic genes and whose expression is driven by the *Pz* and *Pa* promoters, respectively, and the *paaXY* operon expressing the *paaX* regulatory gene from the *Px* promoter (2). The absence of the *paa* genes in *E. coli* W14, an *E. coli* W mutant strain (3), leads to a $PA^-$ phenotype (2).

Previous work involving *Pa-lacZ* translational fusions revealed that the *Pa* promoter is negatively regulated by the *paaX* gene product since the absence of the latter caused a constitutive expression of the reporter fusion (2). The PaaX repressor is a 316-amino acid protein that shows 41.4% amino acid sequence identity with its ortholog PhaN (recently renamed as PaaN (4)) from the PA degradation pathway of *P. putida* U (1) and contains a stretch of 25 residues at amino acids 39–64 that shares similarity with the helix-turn-helix motif for DNA recognition and binding of transcriptional repressors from the GntR family such as FadR (5) and GntR (6). It is worth mentioning that whereas activator proteins are very common in transcriptional regulation of aromatic catabolic operons, only a few repressors have been described so far in biodegradation of aromatic compounds, namely HpaR (HpcR) for the catabolism of homoprotocatechuic acid in *E. coli* (7, 8), CymR for the catabolism of *p*-cymene in *P. putida* F1 (9), and AphS for the catabolism of phenol in *Comamonas testosteroni* TA441 (10).

The repressor effect of PaaX on the *Pa* promoter in *E. coli* W14(*lacZ*$^-$) cells, containing the *Pa-lacZ* fusion into the chromosome and the *paaX* gene in a plasmid, could not be alleviated by growing the cells in the presence of PA, suggesting that this aromatic compound is not the inducer of the pathway (2). However, the simultaneous expression of genes *paaX* and *paaK*, the latter encoding the PA-CoA ligase that catalyzes the activation of PA to PA-CoA, allowed activity of the *Pa* promoter when the cells were grown in the presence of PA, suggesting that PA-CoA rather than PA is the true inducer of the pathway (2).

Due to the unusual catabolic and regulatory features mentioned above, the PA biodegradation pathway becomes a very interesting model of an aerobic hybrid route for the catabolism of aromatic compounds. In this work, we have performed both *in vivo* and *in vitro* experiments to investigate the regulation of gene expression of the *paaZ* and *paaABCDEFGHIJK* catabolic operons from *E. coli*. Superimposed in the specific PaaX-mediated repression, two global regulators, the cAMP receptor protein (CRP) and the integration host factor protein (IHF), act as activators of the gene expression driven by *Pa* and *Pz* promoters.

* This work was supported in part by Grants AMB97-063-C02-02 from the Comisión Interministerial de Ciencia y Tecnología and 06M/029/96 from the Comunidad Autónoma de Madrid, Spain. The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

‡ Recipient of a predoctoral fellowship from the Plan Nacional de Formación de Personal Investigador-MEC. Present address: Dept. of Microbiology, the University of Iowa, Iowa City, IA 52242.

§ To whom correspondence should be addressed: Dept. of Molecular Microbiology, Centro de Investigaciones Biológicas, Consejo Superior de Investigaciones Científicas, Velázquez, 144, 28006 Madrid, Spain. Tel.: 34-91-5611800; Fax: 34-91-5627518; E-mail: ediaz@cib.csic.es.

[1] The abbreviations used are: PA, phenylacetic acid; bp, base pair(s); CRP, cAMP receptor protein; CoA, coenzyme A; IHF, integration host factor; LH, left-half; OR, operator region; PA-CoA, phenylacetyl-coenzyme A; PCR, polymerase chain reaction; RH, right-half.

TABLE I
*Bacteria and plasmids used in this study*

| Strain or plasmid | Relevant genotype and characteristic(s) | Ref. or origin |
|---|---|---|
| *E. coli* K-12 | | |
| DH5α | *endA1 hsdR17 supE44 thi-1 recA1 gyrA*(Nal$^r$) *relA1* Δ(*argF-lac)U169 deoR* ϕ80*dlac*Δ(*lacZ*)*M15* | 11 |
| S17–1λpir | Tp$^r$ Sm$^r$ *recA thi hsdR*$^-$*M*$^+$ *RP4::2-Tc::Mu::Km Tn7* λpir phage lysogen | 12 |
| CC118λpir | Δ(*ara-leu) araD* Δ*lacX74 galE galK phoA thi-1 rpsE* (Sp$^r$) *rpoB*(Rif$^r$) *argE*(am) *recA1* λpir phage lysogen | 12 |
| S90C | Δ(*lac, pro) rpsL* (Sm$^r$) | 13 |
| DPB101 | S90C *himD451*::mini-*tet* | 13 |
| DPB102 | S90C *himA452*::mini-*tet* | 13 |
| MC4100 | *araD319* Δ(*argF-lac)U169 rpsL150* (Sm$^r$) *relA1flbB5301 deoC1 ptsF25 rbsR* | 14 |
| SBS688 | MC4100 Δ*crp* | 14 |
| SH210 | Hfr (PO2A), Δ(*argF-lac*)169, *zai-736*::Tn*10*(Tc$^r$) | 15 |
| AFMC | MC4100 spontaneous rifampicin-resistant mutant (Rif$^r$) | This work |
| AFSB | SBS688 spontaneous rifampicin-resistant mutant (Rif$^r$) | This work |
| S90CRif | S90C spontaneous rifampicin-resistant mutant (Rif$^r$) | This work |
| DPB101Rif | DPB101 spontaneous rifampicin-resistant mutant (Rif$^r$) | This work |
| DPB102Rif | DPB102 spontaneous rifampicin-resistant mutant (Rif$^r$) | This work |
| S90CPA | S90CRif with chromosomal insertion mini-Tn*5*Km2 *Pa-lacZ* | This work |
| S90CPZ | S90CRif with chromosomal insertion mini-Tn5Km2 *Pz-lacZ* | This work |
| DPB101PA | DPB101Rif with chromosomal insertion mini-Tn5Km2 *Pa-lacZ* | This work |
| DPB101PZ | DPB101Rif with chromosomal insertion mini-Tn5Km2 *Pz-lacZ* | This work |
| DPB102PA | DPB102Rif with chromosomal insertion mini-Tn5Km2 *Pa-lacZ* | This work |
| DPB102PZ | DPB102Rif with chromosomal insertion mini-Tn5Km2 *Pz-lacZ* | This work |
| AFMCPA | AFMC with chromosomal insertion mini-Tn*5*Km2 *Pa-lacZ* | This work |
| AFMCPZ | AFMC with chromosomal insertion mini-Tn*5*Km2 *Pz-lacZ* | This work |
| AFSBPA | AFSB with chromosomal insertion mini-Tn*5*Km2 *Pa-lacZ* | This work |
| AFSBPZ | AFSB with chromosomal insertion mini-Tn*5*Km2 *Pz-lacZ* | This work |
| *E. coli* W | | |
| W14 | W Δ*paa* | 3 |
| W14Rif | W14 spontaneous rifampicin-resistant mutant (Rif$^r$) | 2 |
| AF15 | W14Rif Δ*lac* | This work |
| AF15PZ | AF15 with chromosomal insertion mini-Tn*5*Km2 *Pz-lacZ* | This work |
| Plasmids | | |
| pUTmini-Tn*5*Km2 | Ap$^r$ Km$^r$; R6K*oriV* RP4*oriT*, mini-Tn*5* Km2 transposon delivery plasmid | 12 |
| pSJ3 | Apr$^r$; *'lacZ* promoter probe vector, *lacZ* fusion flanked by *Not*I sites | 2 |
| pAFPZ | Ap$^r$; pSJ3 containing a 446-bp *Bbu*I-*Bam*HI DNA fragment to produce the *Pz-lacZ* fusion | 2 |
| pAFPZT | Ap$^r$ Km$^r$; pUTmini-Tn*5* Km2 containing the 4.4-kb[a] *Not*I *Pz-lacZ* fragment from pAFPZ | This work |
| pAFPA1T | Ap$^r$ Km$^r$; pUTmini-Tn*5* Km2 containing the 4.4-kb *Not*I *Pa-lacZ* fragment | 2 |
| pCK01 | Cm$^r$; low copy number cloning vector (pSC101 derivative), polylinker flanked by *Not*I sites | 2 |
| pAAD | Cm$^r$; pCK01 containing a 15.5-kb DNA fragment carrying the *paa* cluster | 2 |
| pAFK5 | Ap$^r$; high copy number pUC19 derivative containing the *paaK* gene under *Plac* | 2 |
| pUC18 | Ap$^r$; high copy number cloning vector | 11 |
| pAFX | Ap$^r$; pUC18 expressing the *paaX* gene under *Plac* | 2 |
| pAFX2 | Cm$^r$; pCK01 expressing the *paaX* gene under *Plac* | 2 |
| pAFPA2 | Ap$^r$; pSJ3 containing a 255-bp *Xba*I-*Bbu*I DNA fragment to produce the *Pa-lacZ* fusion | 2 |
| pAFPA2-M1 | Ap$^r$; pSJ3 containing a 255-bp *Xba*I-*Bbu*I DNA fragment to produce the *PaM1-lacZ* fusion | This work |
| pAFPA2-M5 | Ap$^r$; pSJ3 containing a 255-bp *Xba*I-*Bbu*I DNA fragment to produce the *PaM5-lacZ* fusion | This work |
| pAFPA2-M9 | Ap$^r$; pSJ3 containing a 255-bp *Xba*I-*Bbu*I DNA fragment to produce the *PaM9-lacZ* fusion | This work |
| pAFPA2-M16 | Ap$^r$; pSJ3 containing a 255-bp *Xba*I-*Bbu*I DNA fragment to produce the *PaM16-lacZ* fusion | This work |
| pAFPA2-M18 | Ap$^r$; pSJ3 containing a 255-bp *Xba*I-*Bbu*I DNA fragment to produce the *PaM18-lacZ* fusion | This work |
| pAFPA2-M22 | Ap$^r$; pSJ3 containing a 255-bp *Xba*I-*Bbu*I DNA fragment to produce the *PaM22-lacZ* fusion | This work |
| pAFPA2-M25 | Ap$^r$; pSJ3 containing a 255-bp *Xba*I-*Bbu*I DNA fragment to produce the *PaM25-lacZ* fusion | This work |
| pAFPA2-M26 | Ap$^r$; pSJ3 containing a 255-bp *Xba*I-*Bbu*I DNA fragment to produce the *PaM26-lacZ* fusion | This work |

[a] kb, kilobase pair.

## EXPERIMENTAL PROCEDURES

*Bacterial Strains, Plasmids, and Growth Conditions*—The *E. coli* strains and plasmids used in this work are listed in Table I. The *E. coli* AF15 strain was obtained by transferring the (*argF-lac)U169* deletion of *E. coli* SH210 to *E. coli* W14Rif through biparental mating (12) and selecting for a clone resistant to tetracycline and rifampicin. Unless otherwise stated, bacteria were grown in Luria-Bertani (LB) medium (11) at 37 °C. Growth in M63 minimal medium (16) was achieved at 30 °C using the corresponding necessary nutritional supplements and 20 mM glycerol or 10 mM glucose as carbon source. When required, 1 mM PA was added to the M63 minimal medium. Where appropriate, antibiotics were added at the following concentrations: ampicillin (100 μg/ml), chloramphenicol (35 μg/ml), tetracycline (10 μg/ml), kanamycin (50 μg/ml), and rifampicin (50 μg/ml).

*DNA Manipulations and Sequencing*—Plasmid DNA was prepared by the rapid alkaline lysis method (11). Transformation of *E. coli* was carried out using the RbCl method (11). DNA manipulations and other molecular biology techniques were essentially as described (11). DNA fragments were purified using the Geneclean Kit (Bio 101, Inc.). Oligonucleotides were synthesized on an Oligo-1000M nucleotide synthesizer (Beckman Instruments). Nucleotide sequences were determined directly from plasmids by using the dideoxy chain termination method (17). Standard protocols of the manufacturer for *Taq* DNA polymerase-initiated cycle sequencing reactions with fluorescently labeled dideoxynucleotide terminators (Applied Biosystems Inc.) were used. The sequencing reactions were analyzed using a model 377 automated DNA sequencer (Applied Biosystems Inc.).

*Mutagenesis of the PaaX Binding Motif*—Base substitutions in the PaaX binding motif were introduced through recombinant PCR using the plasmid pAAD as template. To generate mutations in the left-half site of the PaaX binding motif, a first PCR reaction was performed using two pairs of primers as follows: PA5–1 (5′-CAATCTCGGAATGCG-CATG-3′, which encodes amino acids 38–44 of PaaA and hybridizes with the non-coding strand of the *paaA* gene between nucleotides 2885 and 2867 of the *paa* gene cluster (2), underlined are the bases that reconstruct a *Bbu*I restriction site) and the degenerated MUT-PA5 (5′-CTTTCATAAAACAATGWKMTTCGTGTTTTTAATT-3′, which hybridizes between nucleotides 2688 and 2721 enclosing the transcription start sites of *Pa* (2), see also Fig. 4, where K is G or T, M is A or C, and W is A or T); and PAP (5′-GGTCTAGAGTTATCAAAATAGAGTGCG-3′, located between the *Pz* and *Pa* promoters, nucleotides 2611–2631 of the *paa* gene cluster (2) (see Fig. 4), where the engineered *Xba*I site is underlined) and the degenerated MUT-PA3 (5′-AATTAAAAACAC-GAAKMWCATTGTTTTATGAAAG-3′, which is complementary to the MUT-PA5 oligonucleotide). The resulting two PCR products were mixed and subjected to a further PCR reaction using PAP and PA5–1 as

primers. Similarly, to generate mutations in the right-half site of the PaaX binding motif, a first PCR reaction was performed using two pairs of primers: PA5–1 and the degenerated MUT-PA52 (5′-CATAAAACAATGTGATTCGWSWWTTTAATTAATTCACG-3′, which hybridizes between nucleotides 2692 and 2729 enclosing the transcription start sites of *Pa* (2), see also Fig. 4; where S means C or G); and PAP and the degenerated MUT-PA32 (5′-CGTGAATTAATTAAAWWSWCGAATCACATTGTTTTATG-3′, which is complementary to the MUT-PA52 oligonucleotide). The resulting two PCR products were mixed and subjected to a further PCR reaction using PAP and PA5–1 as primers. The amplified 272-bp fragments were digested with *Xba*I and *Bbu*I and cloned as 255-bp fragments into the double-digested *Xba*I+*Bbu*I pSJ3 promoter probe vector, giving rise to the pAFPA2-M plasmids (Table I), which were further sequenced to confirm the nucleotide sequence of the cloned mutant fragments.

*Construction of E. coli Strains Harboring Chromosomal Insertions of the Pa-lacZ and Pz-lacZ Translational Fusions*—By means of RP4-mediated mobilization, plasmids pAFPA1T and pAFPZT, which contain mini-Tn*5*Km2 hybrid transposons expressing *Pa-lacZ* and *Pz-lacZ* fusions, respectively, were transferred from *E. coli* S17–1λpir into different rifampicin-resistant *E. coli* recipient strains, *i.e.* AF15, S90CRif, DPB101Rif, DPB102Rif, AFMC, and AFSB, as described previously (12). Exconjugants containing the *lacZ* translational fusions stably inserted into the chromosome were selected for the transposon marker, kanamycin, on rifampicin-containing LB medium. The resulting strains, AF15PZ, S90CPA, S90CPZ, DPB101PA, DPB101PZ, DPB102PA, DPB102PZ, AFMCPA, AFMCPZ, AFSBPA, and AFSBPZ, and their relevant genotype are indicated in Table I.

*Production of the PaaX Repressor*—The *paaX* gene was expressed from the *Plac* promoter in the high copy number pAFX plasmid and in the low copy number pAFX2 plasmid. SDS-polyacrylamide gel electrophoresis analyses of crude lysates from *E. coli* cells harboring plasmids pAFX or pAFX2 did not show the presence of an intense band corresponding to the PaaX protein when they were compared with crude lysates from cells harboring the parental plasmids, thus indicating that the *paaX* gene was not overexpressed in these recombinant plasmids.

To prepare crude extracts containing the PaaX protein, $PaaX^+$ extracts, *E. coli* W14 (pAFX) cells were grown in ampicillin-containing LB medium to an $A_{600}$ of about 1. Cell cultures were then centrifuged ($3,000 \times g$, 10 min at 20 °C), and cells were washed and resuspended in 0.05 volumes of 20 mM Tris-HCl buffer, pH 7.5, containing 10% glycerol, 2 mM β-mercaptoethanol, and 50 mM KCl, prior to disruption by passage through a French press (Aminco Corp.) operated at a pressure of 20,000 pounds/square inch. The cell debris was removed by centrifugation at $26,000 \times g$ for 30 min at 4 °C. The clear supernatant fluid was decanted and used as crude cell extract. The $PaaX^-$ extracts from *E. coli* W14 (pUC18) cells were prepared in a similar manner. Protein concentration was determined by the method of Bradford (18) using bovine serum albumin as standard.

*Synthesis of DNA Fragments Covering the paaZ-paaA Intergenic Region*—The target DNA fragments used as probes were generated by PCR using plasmid pAAD as template. To prepare the Pz-Pa fragment (483-bp), primers PZ5 (5′-GGGGTGAATCAAACGGCTACG-3′, which encodes amino acids 18–23 of PaaZ and hybridizes with the non-coding strand of the *paaZ* gene between nucleotides 2402 and 2420 of the *paa* gene cluster (2)) and PA5–1 were used. To prepare the Pa fragment (274-bp), primers PAP and PA5–1 were used. To prepare the Pz fragment (242-bp), primers SG-1 (5′-CTGTGACAGATTTCGCACTC-3′, which hybridizes between nucleotides 2644 and 2625 of the *paaZ-paaA* intergenic region (2), see also Fig. 4) and PZ5 were used.

The 217-bp DNA fragments harboring mutations in the PaaX binding motif were PCR-amplified from the pAFPA2-M series of plasmids (Table I) using primers PAP and PA5–4 (5′-CGGGCATCCAGTCCTGTGGCTCG-3′, which encodes amino acids 18–25 of PaaA and hybridizes with the non-coding strand of the *paaA* gene between nucleotides 2828 and 2806 (2)). The PAP and PA5–4 primers were also used to amplify the wild-type *Pa* promoter from the pAFPA2 plasmid.

*Gel Retardation Assays*—DNA fragments used as probes were labeled at their 5′-end with phage T4 polynucleotide kinase (Amersham Pharmacia Biotech) and [$\gamma$-$^{32}$P]ATP (Amersham Pharmacia Biotech). The reaction mixtures contained 20 mM Tris-HCl, pH 7.5, 10% glycerol, 2 mM β-mercaptoethanol, 50 mM KCl, 0.1 nM DNA probe, 50 μg/ml bovine serum albumin, 50 μg/ml salmon sperm (competitor) DNA, and cell extract or purified IHF or CRP in a 20-μl final volume. Purified IHF and CRP were kindly provided by S. Goodman and A. Kolb, respectively. After incubation for 20 min at 30 °C, mixtures were fractionated by electrophoresis in 4% polyacrylamide gels buffered with 0.5× TBE (45 mM Tris borate, 1 mM EDTA). The gels were dried onto Whatman 3MM paper and exposed to Hyperfilm MP (Amersham Pharmacia Biotech). For quantitative definition of the PaaX-DNA interactions, autoradiograms were scanned in a computing densitometer, and the intensity of the bound and unbound DNA bands was measured using the ImageQuant software (Molecular Dynamics). The apparent dissociation constant ($K_{d(app)}$) of PaaX for the probe DNA was defined as the amount of total protein per ml of reaction mixture at which half of the labeled DNA remained unbound.

*DNase I Footprinting Assays*—DNA fragments to be used as probes were singly 5′-end-labeled by using a labeled primer during the PCR-amplification reaction. The desired primer (50 pmol) was 5′-end-labeled with [$\gamma$-$^{32}$P]ATP using T4 polynucleotide kinase. Unincorporated nucleotides were removed in a G-25 Sephadex column (Amersham Pharmacia Biotech), and the labeled primer was eluted with 100 μl of water. To perform the PCR reaction, 25 pmol of both the labeled and unlabeled primers were used. The 5′-end-labeled PCR product was separated in a 2% agarose gel and purified through the Geneclean DNA purification kit (Bio 101). For DNase I footprinting assays, the reaction mixture contained 20 mM Tris-HCl, pH 7.5, 10% glycerol, 2 mM β-mercaptoethanol, 50 mM KCl, 0.1 nM DNA probe, 50 μg/ml bovine serum albumin, 50 μg/ml salmon sperm (competitor) DNA, and cell extract in a 100-μl final volume. This mixture was incubated for 20 min at 30 °C, after which 2 μl (0.005 units) of DNase I (Sigma) (prepared at 1 ng/μl in 125 mM $CaCl_2$, 50 mM $MgCl_2$) was added, and the incubation was continued for 2 min at 37 °C. After phenol/chloroform extraction, ammonium acetate was added to the clear supernatant to a final concentration of 0.2 M, and DNA fragments were precipitated with ethanol absolute, washed with 70% ethanol, and directly resuspended in 5 μl of 90% (v/v) formamide-loading gel buffer (10 mM Tris-HCl, pH 8.0, 20 mM EDTA, pH 8.0, 0.05% (w/v) bromphenol blue, 0.05% (w/v) xylene cyanol). Samples were then denatured at 95 °C for 2 min and fractionated in a 8% polyacrylamide-urea gel. A + G Maxam and Gilbert reactions (19) were carried out with the same fragments and loaded in the gels along with the footprinting samples. The gels were dried onto Whatman 3MM paper and exposed to Hyperfilm MP (Amersham Pharmacia Biotech).

*β-Galactosidase Assays*—β-Galactosidase activities were measured with permeabilized cells as described by Miller (16).

## RESULTS

*The PaaX Repressor Regulates the Pz Promoter*—We have previously shown that the *Pa* promoter, which drives the expression of the *paaABCDEFGHIJK* catabolic operon, is repressed by the PaaX regulator (2). To check whether the expression of the divergently transcribed *Pz* promoter, which drives the expression of the *paaZ* catabolic gene, was also regulated by PaaX, we have engineered a reporter *Pz-lacZ* fusion within a mini-Tn*5* vector. The resulting construction, pAFPZT, was used to deliver by transposition the *Pz-lacZ* translational fusion into the chromosome of *E. coli* AF15, a Δ*lac* derivative of *E. coli* W14Rif (Δ*paa*), giving rise to the reporter strain *E. coli* AF15PZ. The β-galactosidase assays of permeabilized *E. coli* AF15PZ cells harboring control plasmid pUC19 or the plasmid pAFK5, a pUC19 derivative that expresses the *paaK* gene encoding the PA-CoA ligase, revealed a similar and constitutive expression of the *Pz-lacZ* fusion (Table II). When the *paaX* gene was expressed in *trans* from the pAFX2 plasmid in *E. coli* AF15PZ cells grown in glycerol-containing minimal medium in the presence or in the absence of 1 mM PA, no β-galactosidase activity was detected (Table II). However, when the *paaX* and *paaK* genes were simultaneously expressed in the reporter strain, β-galactosidase activity was observed when the cells were grown in the presence of 1 mM PA, and no activity was detected in the absence of this aromatic compound (Table II). These data, therefore, indicate that PaaX behaves as a transcriptional repressor of *Pz* and suggest that the product of the reaction catalyzed by PaaK, *i.e.* PA-CoA, is the true inducer of this promoter.

*Binding of PaaX to the Pa and Pz Promoter Regions*—Cell-free extracts from *E. coli* W14 (pAFX), a strain that expresses the *paaX* gene under the *Plac* promoter, were subjected to gel retardation analysis using as probe a 483-bp DNA fragment carrying the *Pz* and *Pa* promoters (Pz-Pa probe). As shown in

TABLE II
*Regulation of expression from the Pz promoter by PaaX*

*E. coli* AF15PZ (*Pz-lacZ*) cells containing different plasmids were grown in glycerol-containing minimal medium in the absence (uninduced) or in the presence (induced) of 1 mM PA, until the cultures reached an $A_{600}$ of 1.0. β-Galactosidase activities were measured with permeabilized cells as described under "Experimental Procedures." Results of one experiment are shown; values were reproducible in three separate experiments.

| Plasmid(s) | Relevant gene(s) | β-Galactosidase activity (Miller units) | | |
|---|---|---|---|---|
| | | Uninduced | Induced (1 mM PA) | -Fold induction |
| pUC19 | | 940 | 950 | 1.0 |
| pAFK5 | *paaK* | 900 | 916 | 1.0 |
| pAFX2 | *paaX* | BD[a] | BD | |
| pAFK5/pAFX2 | *paaK*/*paaX* | BD | 500 | >500 |

[a] BD, below detection limits.

Fig. 1, whereas extracts containing PaaX were able to retard the migration of the Pz-Pa probe in a protein concentration-dependent manner, control extracts prepared from *E. coli* W14 (pUC18) cells, did not. Moreover, whereas a 6000-fold excess amount of an unrelated DNA such as that of salmon sperm did not affect the binding of PaaX to the target DNA, migration of the labeled fragment was not retarded when a 10-fold excess amount of non-labeled Pz-Pa fragment was added to the assay (data not shown), indicating that binding of the PaaX repressor to the *Pz-Pa* promoter region was specific. As the amount of PaaX protein increased, two distinct complexes, complex 1 (higher mobility) and 2 (lower mobility), were observed (Fig. 1), a result that might suggest that PaaX binds with different affinities to two sites in the fragment used as probe. To confirm the existence of two different PaaX-binding sites, gel retardation assays were performed using the 242-bp Pz fragment (from position −146 to +97 relative to the transcription start site of *Pz* promoter) and the 274-bp Pa fragment (from position −87 to +187 relative to the major transcription start site of *Pa* promoter), as probes. As shown in Fig. 1, one shifted band was found with each individual probe fragment. Since the deduced $K_{d(app)}$ of PaaX (see under "Experimental Procedures") for the Pa and Pz probes, 7 and 150 μg protein/ml, respectively, were similar to that for complex 1 and complex 2 with the Pz-Pa probe, there is no cooperativity in the binding of PaaX to the *Pa* and *Pz* promoters. Therefore, these data taken together revealed the specific and independent binding of PaaX to the *Pz* and *Pa* promoters.

*PA-CoA Specifically Inhibits Binding of PaaX to Its Cognate Promoter Region*—We have shown that the *lacZ* expression driven by the *Pa* (2) and *Pz* (see above) promoters in cells containing the *paaX* gene in a plasmid requires the simultaneous expression of the *paaK* gene encoding the PA-CoA ligase activity, as well as the presence of PA in the culture medium. This has led to the suggestion that the true inducer of the *paa* catabolic cluster is PA-CoA rather than PA. Nevertheless, since acyl-CoA ligases have been shown to play a role in the cellular uptake of some substrates such as long chain fatty acids (20), we could not exclude the possibility that *E. coli* cells lacking the *paaK* gene could have some deficiency in PA uptake that would lead to noninducible *paa* cluster expression. We therefore tested by gel retardation assays the ability of PA and PA-CoA to inhibit binding of PaaX to the *Pa* promoter region *in vitro* (Fig. 2). Whereas 30 μM PA-CoA reduced the amount of bound DNA (Pa·PaaX complex) to 50%, PA did not affect the PaaX-DNA binding (Fig. 2), even when this compound was added up to a final concentration of 2.5 mM (data not shown). A close analogue of PA-CoA such as benzoyl-CoA had no effect on PaaX-DNA binding when added to the gel retardation assay up to a final concentration of 2.5 mM (data not shown). Other CoA derivatives such as acetyl-CoA and free CoA were also unable to abolish binding of PaaX to the *Pa* promoter (data not shown), thus confirming PA-CoA as the specific inducer molecule of the *paa* catabolic genes.

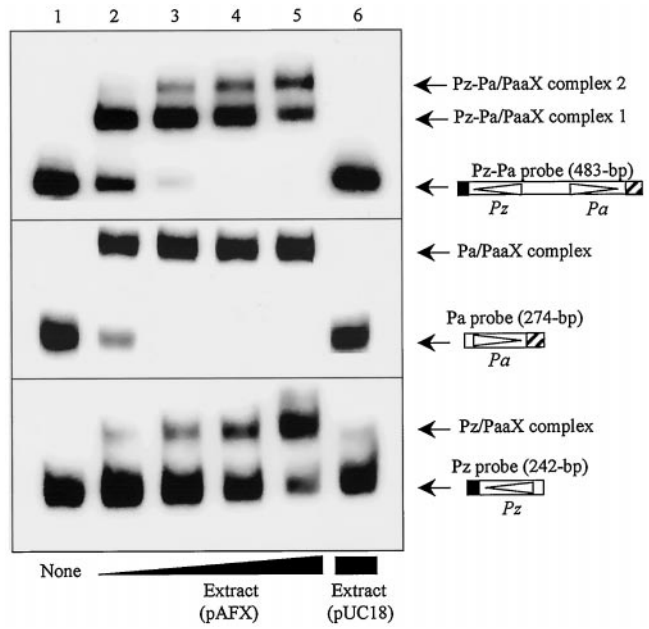


FIG. 1. **Gel retardation analyses of PaaX binding to the *paaZ-paaA* promoter region.** The probe DNAs used (Pz-Pa, Pa, and Pz) are shown schematically on the *right*, with the *Pa* and *Pz* promoters marked with *open arrowheads* and the first 44 and 23 amino acids from the *paaA* and *paaZ* gene products represented by *striped* and *black blocks*, respectively. Cell extract and gel retardation analyses were performed as described under "Experimental Procedures." *Lanes 2–5,* contained 0.3, 1.0, 2.0, and 4.0 μg of total protein of $PaaX^+$ extracts obtained from cells bearing plasmid pAFX (*paaX*), respectively. *Lanes 1* and *6* contained no extract and 4.0 μg of total protein of $PaaX^-$ extracts obtained from cells bearing control plasmid pUC18, respectively. The DNA·PaaX complexes are indicated.

*Identification of the PaaX Binding Regions*—To localize the PaaX binding region within the *Pa* and *Pz* promoters, DNase I footprinting experiments were performed using the same target DNA fragments as those used in the gel retardation assays reported above. As shown in Fig. 3, an extended binding region was detected in each promoter. The footprinting assay with the non-coding strand of the Pa fragment (Fig. 3*B*) revealed a protected region, operator region A ($OR_A$), between positions −1 and +51 with respect to the major transcription start site of *paaA* that partially overlaps the ribosome-binding site of the gene (Fig. 4). On the other hand, the non-coding strand of the Pz fragment was protected by PaaX from DNase I digestion between positions −30 and +17, operator region Z ($OR_Z$), spanning the potential −10 box of the promoter (Figs. 3*C* and 4). Similar patterns of DNase I digestion for each of the coding strands were observed (Fig. 3, *A* and *D*). The DNase I-hypersensitive sites observed in $OR_A$ and $OR_Z$ were spaced at approximately 10 nucleotide intervals, corresponding to about one helix turn (Figs. 3 and 4), which suggests binding of PaaX to one side of the double helix.

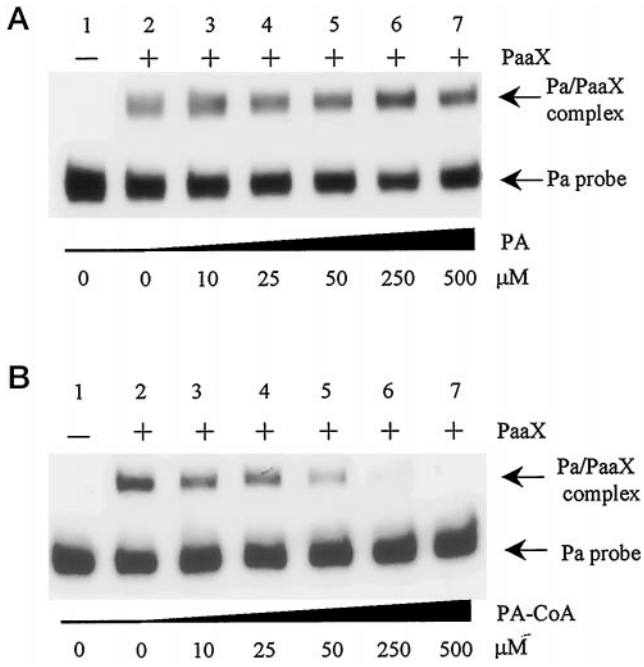


FIG. 2. **Effect of PA and PA-CoA on PaaX-specific binding to its target DNA.** Gel retardation analyses were performed as described under "Experimental Procedures" using the Pa fragment as probe. − and + indicate that no extract (*lane 1*) and 0.2 μg of total protein of $PaaX^+$ extracts (*lanes 2–7*) were added to the reaction mixtures, respectively. PA (*A*) and PA-CoA (*B*) were added to the reaction mixtures at the indicated final concentration. Unbound Pa probe and the Pa·PaaX complex are shown by *arrows*.

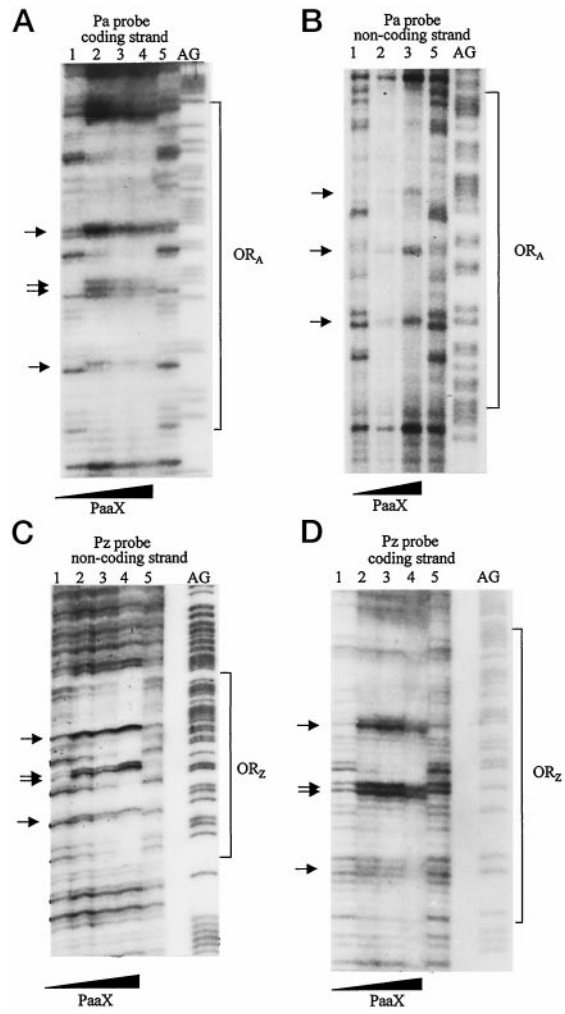


FIG. 3. **DNase I footprinting analyses of the interaction of PaaX with the *paaA* and *paaZ* promoter regions.** The DNase I footprinting experiments were carried out using as probes 5′-end-labeled non-coding and coding strands of the Pa (274-bp) (*A* and *B*) and Pz (242-bp) (*C* and *D*) fragments, as indicated under "Experimental Procedures." *Lanes 1–4* contained 0, 3, 6, and 15 μg of total protein of $PaaX^+$ extracts, respectively. *Lane 5* contained 15 μg of total protein of $PaaX^-$ extracts. AG lane corresponds to the A + G Maxam and Gilbert sequencing reaction. Protected regions, $OR_A$ and $OR_Z$, are shown in the autoradiograms by *brackets*. DNase I-hypersensitive sites are indicated by *arrows*.

*Characterization of a Conserved PaaX Binding Motif*—Alignment of the PaaX binding sequences identified by DNase I footprinting analyses revealed that each of the operator regions, $OR_A$ in promoter *Pa* and $OR_Z$ in promoter *Pz*, contained a conserved 15-bp imperfect palindromic motif (Fig. 4). Whereas the conserved motif in $OR_A$, AATGTGATTCGTGTT, is located between positions +3 and +17, that of $OR_Z$, TTTATGATTCGCGAT, is located between positions −30 and −16 of the non-coding strands (Fig. 4). A consensus palindromic sequence was deduced to be WWTRTGAT*TCGYGYWT* (R is A or G; W is A or T; and Y is C or T), with its pseudo-dyad axis through the central T base (underlined) which defines a left-half (LH) and a right-half (RH) (in *italics*) region. Since sequences with dyad symmetry within the operator regions are usually the binding sites of bacterial transcriptional regulators (21), we predicted that the described conserved motifs might be the binding sites of PaaX within $OR_A$ and $OR_Z$.

To check the involvement of the conserved motif in the recognition and binding of PaaX to $OR_A$, base substitutions in the target motif were generated by recombinant PCR (Fig. 5*A*), and the amplified wild-type and mutant Pa fragments were ligated to the *′lacZ* gene of the promoter-probe vector pSJ3 as indicated under "Experimental Procedures." Since the β-galactosidase levels of permeabilized *E. coli* AF15 cells expressing the wild-type *Pa-lacZ* translational fusion were similar to that of cells expressing the different mutant *Pa-lacZ* fusions (Table III), the possibility that mutations in $OR_A$ could negatively influence the maximum activity of the *Pa* promoter can be ruled out. The binding of PaaX to the different promoter fragments was checked *in vivo* by assaying β-galactosidase activity in *E. coli* AF15 cells expressing simultaneously *paaX* and the different *lacZ* translational fusions. As shown in Table III, base substitutions in the LH region of the conserved motif, *i.e.* mutant promoters *PaM1, PaM5, PaM9*, and *PaM16* (Fig. 5*A*), severely impaired the PaaX-mediated repression of the corresponding *lacZ* fusions as compared with the wild-type *Pa-lacZ* fusion. As expected, the reduction in repression was higher with mutant promoters *PaM1* and *PaM9,* both containing two base substitutions, than with mutant promoters *PaM5* and *PaM16* bearing only one base substitution (Table III). Interestingly, mutations in the RH region of the conserved motif, *i.e.* mutant promoters *PaM18, PaM25*, and *PaM26* (Fig. 5*A*), did not cause a reduction in PaaX-mediated repression as high as that observed with base substitutions in the LH region. Thus, deletion of the conserved guanine at position 11 of the mutant promoter *PaM18* (Fig. 5*A*) led to a 3-fold reduction in repression compared with that observed with the wild-type *Pa* promoter, and the two base substitutions in mutant promoters *PaM25* and *PaM26* (Fig. 5*A*) maintained the strong repression effect of PaaX on the corresponding *lacZ* translational fusions (Table III).

The effects of the base substitutions on binding of PaaX to $OR_A$ were also tested *in vitro* by using the wild-type and the different mutant Pa promoter fragments as probes in gel retardation assays. Whereas mutations in the LH region of the


FIG. 4. **Nucleotide sequence of the *paaZ-paaA* intergenic region.** The translation initiation codon for the *paaZ* and *paaA* genes (represented by *arrows*) is shown by *lowercase type* in the corresponding non-coding strands. *M* and *V*, is the standard *one-letter code* abbreviation for the first amino acid of PaaZ and PaaA, respectively. The ribosome-binding site (*RBS*) is shown in *italics*. The transcription start sites in the *Pz* and *Pa* promoters (2) are *double-underlined,* and the direction of transcription is indicated by *bent arrows*. The major transcription start site in *Pa* is shown in *boldface type*. The inferred −10 and −35 regions of each promoter are *underlined*. The PaaX-mediated protection of the coding and non-coding strands from digestion by DNase I is indicated with *boldface letters*. The PaaX binding motif is *boxed*, and its dyad symmetry indicated by *convergent arrows*. DNase I-hypersensitive sites in each strand are marked by *asterisks*. Putative IHF and CRP binding sequences are shown by *brackets*. The 5′-ends of the PAP and SG-1 primers used to generate the Pa and Pz fragments (see the "Experimental Procedures"), respectively, are indicated by *bent arrows* with *broken lines*.

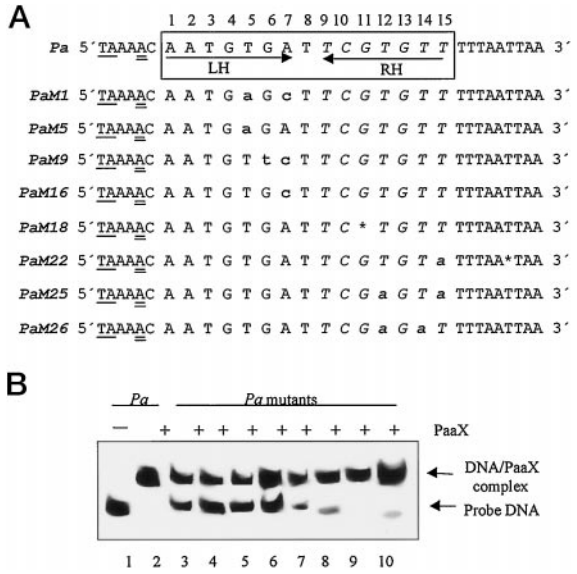


FIG. 5. **Base substitution analyses within the PaaX binding motif.** *A*, alignment of the sequences of wild-type and different mutant *Pa* promoters. The PaaX binding motif is *boxed*, with its LH and RH (in *italics*) regions indicated by *convergent arrows. Numbers* indicate the position of the nucleotides within the 15-bp motif. *Underlines* and *double-underlines* indicate the minor and major transcription start sites, respectively. Base substitutions are shown in *lowercase letters. Asterisks* indicate deleted nucleotides. *B*, binding affinity of PaaX to wild-type and mutant *Pa* promoters. Gel retardation assays were performed as described under "Experimental Procedures." − and + indicate that no extract (*lane 1*) and 0.4 μg of total protein of PaaX$^+$ extract (*lanes 2–10*) were added to the reaction mixtures, respectively. The 217-bp probe DNAs used were generated as reported under "Experimental Procedures," and they are as follows: wild-type Pa (*lanes 1* and *2*), mutant PaM1 (*lane 3*), mutant PaM5 (*lane 4*), mutant PaM9 (*lane 5*), mutant PaM16 (*lane 6*), mutant PaM18 (*lane 7*), mutant PaM22 (*lane 8*), mutant PaM25 (*lane 9*), and mutant PaM26 (*lane 10*). The unbound probe and the DNA·PaaX complexes are indicated by *arrows*.

TABLE III

*PaaX-mediated repression on different Pa-lacZ reporter fusions*

*E. coli* AF15 cells bearing plasmid pAFPA2, which contains the wild-type *Pa-lacZ* fusion, or the different pAFPA2-M plasmids, which contain *lacZ* fusions to the *Pa* mutant promoters indicated in Fig. 5*A*, and the compatible plasmid pAFX2 (*paaX*) (+PaaX) or the control plasmid pCK01 (−PaaX) were grown at 30 °C in glycerol-containing minimal medium to an $A_{600}$ of 1.0. β-Galactosidase activities were measured with permeabilized cells as described under "Experimental Procedures." Results of one experiment are shown; values were reproducible in three separate experiments.

| Plasmid | Relevant genotype | β-Galactosidase activity (Miller units) | | |
|---|---|---|---|---|
| | | −PaaX | +PaaX | -Fold repression |
| pAFPA2 | *Pa-lacZ* | 9133 | 300 | 30.4 |
| pAFPA2-M1 | *PaM1-lacZ* | 9715 | 8205 | 1.1 |
| pAFPA2-M5 | *PaM5-lacZ* | 9625 | 2235 | 4.3 |
| pAFPA2-M9 | *PaM9-lacZ* | 8500 | 4511 | 1.8 |
| pAFPA2-M16 | *PaM16-lacZ* | 10,000 | 2116 | 4.7 |
| pAFPA2-M18 | *PaM18-lacZ* | 9892 | 923 | 10.7 |
| pAFPA2-M22 | *PaM22-lacZ* | 10,000 | 1454 | 6.8 |
| pAFPA2-M25 | *PaM25-lacZ* | 9766 | 474 | 20.6 |
| pAFPA2-M26 | *PaM26-lacZ* | 9077 | 449 | 20.2 |

conserved motif (mutant promoters *PaM1, PaM5, PaM9*, and *PaM16*) significantly reduced the binding affinity of PaaX to the mutant promoters as compared with that for the wild-type *Pa* promoter, mutations in the RH region (mutant promoters *PaM25* and *PaM26*) did not (Fig. 5*B*). The affinity of PaaX for the mutant promoter *PaM18* was lower than that observed for the other two promoters mutated in the RH region, *i.e. PaM25* and *PaM26*, but higher than that for the promoters mutated in the LH region (Fig. 5*B*). These *in vitro* assays, therefore, are in agreement with the *in vivo* results reported above using the translational *lacZ* fusions. Taken together, these data suggest that the conserved motif present in the *Pa* and *Pz* promoters is critical for PaaX repressor binding, with the LH region of the motif being much more important for repressor binding than the RH region.

The mutant promoter *PaM22*, which in addition to a single base substitution at position 15 harbors the deletion of a nucleotide located six bases downstream of the RH region (Fig. 5*A*), led to a 6.8-fold reduction in repression (Table III) and to a significant decrease in PaaX binding affinity (Fig. 5*B*) as compared with the wild-type *Pa* promoter. This suggests that sequences outside the conserved motif may also influence the repressor effect of PaaX on the expression of the *paa* catabolic genes.

*IHF and CRP Behave as Transcriptional Activators of the Pa and Pz Promoters*—Analysis of the *paaZ-paaA* intergenic region revealed the existence of two sequences, TATCAACTAC-TAA and GAGTGCGAAATCTGTCACAGTT, from positions −125 to −113 and −73 to −52 relative to the major transcription start site of *paaA* (Fig. 4), that significantly match the consensus sequences WATCAANNNNTTR (where W is A or T, R is A or G, and N is any of the four bases) and AANTGT-

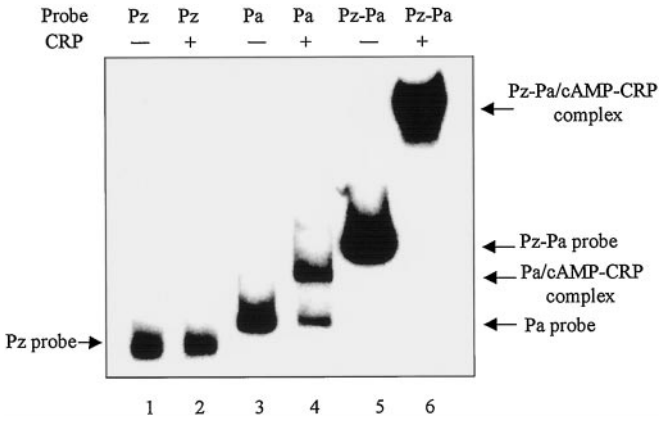


FIG. 6. **CRP binding to the *paaZ-paaA* intergenic region.** Gel retardation analyses were performed as indicated under "Experimental Procedures" but adding 200 $\mu$M cAMP to the reaction mixture and to the electrophoresis buffer. The probe DNAs used were the Pz (*lanes 1* and *2*), Pa (*lanes 3* and *4*), and Pz-Pa (*lanes 5* and *6*) fragments. − and +, indicate the absence and presence of 250 nM purified CRP, respectively. The unbound probe DNAs and the DNA/cAMP·CRP complexes are indicated by *arrows*.

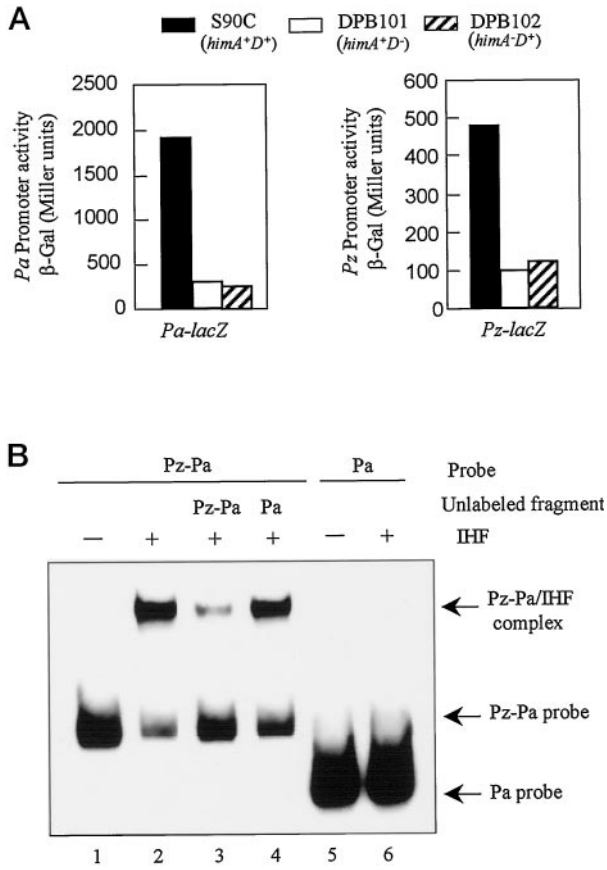


FIG. 7. **Effect of IHF on expression and binding to the *Pa* and *Pz* promoters.** *A*, IHF effect on *Pa* and *Pz* promoter activity *in vivo* in isogenic *E. coli* IHF$^+$ (*himA*$^+$*D*$^+$) and IHF$^-$ (*himA*$^-$*D*$^+$, himA$^+$D$^-$) strains. *E. coli* S90CPA (*Pa-lacZ*), DPB101PA (*Pa-lacZ*), DPB102PA (*Pa-lacZ*), S90CPZ (*Pz-lacZ*), DPB101PZ (*Pz-lacZ*), and DPB102PZ (*Pz-lacZ*) cells were grown in glycerol-containing minimal medium in the presence of 1 mM PA until the cultures reached an $A_{600}$ of 0.7. Values for $\beta$-galactosidase activity in permeabilized cells are shown (in Miller units). Results of one experiment are shown, and values were reproducible in three separate experiments. *B*, IHF binding to the *paaZ-paaA* intergenic region. Gel retardation analyses were performed as indicated under "Experimental Procedures," using as probe DNAs the Pz-Pa (*lanes 1–4*) and Pa (*lanes 5* and *6*) fragments. − and + indicate the absence and presence of 250 nM purified IHF, respectively. Unlabeled Pz-Pa and Pa fragments were added at 10 nM to the reaction mixtures of *lanes 3* and *4*, respectively. The unbound probe DNAs and the Pz-Pa·IHF complex are indicated by *arrows*.

GANNTNNNTCACANTT for binding to the IHF (22) and CRP (23) global regulators (24), respectively. To determine the putative role of CRP and IHF in *paa* regulation, we have used both genetic and biochemical approaches.

Permeabilized cells of *E. coli* AFMCPA and AFMCPZ, two *E. coli* AFMC derivatives which respectively contain a *lacZ* translational fusion with the *Pa* and *Pz* promoters stably inserted into their chromosome, showed $\beta$-galactosidase activities of 2300 and 980 Miller units, respectively, when they were grown in glycerol-containing minimal medium in the presence of PA. However, when these strains were grown in glucose-containing minimal medium in the presence of PA, no $\beta$-galactosidase activity was detected, thus suggesting a repression effect on *Pa* and *Pz* by glucose. To determine whether the observed regulation by glucose was mediated by CRP, the *Pa-lacZ* and *Pz-lacZ* fusions were integrated into the chromosome of an isogenic CRP$^-$ strain, *E. coli* AFSB, as described under "Experimental Procedures." No $\beta$-galactosidase activity was detected in the CRP$^-$ strains AFSBPA (*Pa-lacZ*) and AFSBPZ (*Pz-lacZ*) when the cells were grown in glycerol-containing minimal medium in the presence of PA. Therefore, both the *Pa* and *Pz* promoters seem to be activated by the cAMP-CRP complex, which explains their catabolite repression by glucose reported above. To confirm these genetic experiments, gel retardation assays were performed with purified CRP in the presence of cAMP and three different DNA fragments (Pz-Pa, Pz, and Pa) used as probes. As shown in Fig. 6, whereas a CRP-DNA complex was observed with the Pz-Pa and the Pa probes, no binding of CRP to the Pz probe was detected. These results are also in agreement with the observation that a putative CRP-binding site is located in the *paaZ-paaA* intergenic region but closer to the *Pa* than to the *Pz* transcriptional start site (Fig. 4).

To study the IHF influence on the expression of the *paa* catabolic genes, the *Pa-lacZ* and *Pz-lacZ* translational fusions were stably inserted into the chromosome of isogenic *E. coli* S90CRif (*himA*$^+$*himD*$^+$), DPB101Rif (*himA*$^+$*himD*$^-$), and DPB102Rif (himA$^-$*himD*$^+$) strains (*himA* and *himD* encode the two subunits of the IHF heterodimer) as described under "Experimental Procedures." The resulting strains were grown in glycerol-containing minimal medium in the presence of PA, and the IHF effect on the *lacZ* expression driven by the *Pa* and *Pz* promoters was checked by $\beta$-galactosidase assays. As shown in Fig. 7*A*, there was a significant decrease of *Pa* and *Pz* promoter activity in the IHF$^-$ strains DPB101PA (*Pa-lacZ*), DPB101PZ (*Pz-lacZ*), DPB102PA (*Pa-lacZ*), and DPB102PZ (*Pz-lacZ*), compared with the promoter activity detected in the IHF$^+$ strains S90CPA (*Pa-lacZ*) and S90CPZ (*Pz-lacZ*). To demonstrate whether the IHF stimulation observed *in vivo* from the *Pa* and *Pz* promoters was due to the binding of this regulator to such promoters, gel retardation assays were carried out. Whereas purified IHF was able to bind to the Pz-Pa fragment, no binding to the Pa fragment was detected (Fig. 7*B*). Moreover, the interaction of IHF with the *Pz-Pa* intergenic region was specific since a 100-fold excess of unlabeled Pz-Pa fragment prevented the formation of the IHF-DNA complex but an excess of unlabeled Pa fragment did not.

## DISCUSSION

In this work, we have used both genetic and biochemical approaches to study the regulation of the expression of the divergently transcribed *paaZ* and *paaABCDEFGHIJK* catabolic operons responsible for PA catabolism in *E. coli*. We had shown previously that the PaaX protein behaves as a transcriptional repressor of the *Pa* promoter (2). Here, a genetic analysis based on *lacZ* translational fusions has shown that PaaX acts

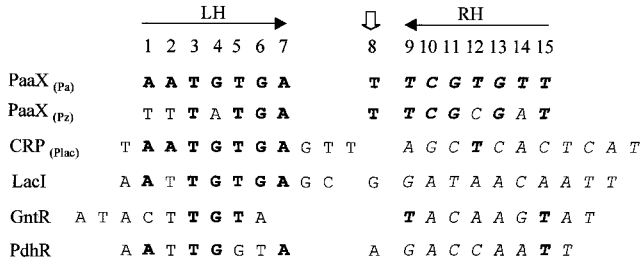


FIG. 8. **Comparison of the PaaX-binding site with other palindromic operator sequences.** Alignment was done with the nucleotide sequences of the binding sites for the PaaX repressor in *Pa* (PaaX(Pa)) and *Pz* (PaaX(Pz)) promoters, the CRP activator in *lac* promoter (CRP-(Plac)) (25), the LacI repressor (26), the GntR repressor (6), and the PdhR repressor (27). The (pseudo) dyad symmetry axis of each sequence is indicated by a *vertical arrow* at the *top*. Nucleotides of the RH region are shown in *italics,* and *numbers* refer to the position of the nucleotides within the 15-bp PaaX binding motif. *Convergent arrows* indicate the LH and RH regions of the PaaX binding motif. The nucleotides identical to that of PaaX(Pa) are shown in *boldface type*.

also as a repressor of the *Pz* promoter (Table II). The results of the gel retardation assays (Fig. 1) indicated that the PaaX repressor binds specifically and independently to each of the *Pa* and *Pz* promoters, although the binding affinity of PaaX for *Pa* was 21-fold higher than that for *Pz*. DNase I footprinting (Fig. 3) revealed that PaaX protected an extended region of about 50-bp that is located immediately downstream of the transcription start sites within the *Pa* promoter ($OR_A$) and that spans the +1 and the −10 region in the *Pz* promoter ($OR_Z$) (Fig. 4). Although it is likely that the effect of the PaaX protein on the *Pa* and *Pz* promoters will involve different mechanisms of repression, validation of this assumption requires further research.

Sequence comparison analysis between $OR_A$ and $OR_Z$ revealed a 15-bp consensus motif, WWTRTGATTCGYGWT, located at the 5′-end of the non-coding strands (Fig. 4), which was shown by point mutation analyses to be indispensable for binding of PaaX to each promoter. The dyad symmetry of the PaaX-binding site and the probability that the repressor-DNA interaction is mediated by a putative helix-turn-helix motif, amino acid residues 39–64 of PaaX (2), that shares similarity to that of transcriptional repressors of the GntR family (5, 6) suggest that the regulator is in an oligomeric form when bound to DNA. The two sequence repeats of the PaaX binding motif are related by a (pseudo) dyad axis through the central T base at position 8 (Fig. 8). Whereas the LH region of the PaaX binding motif is rather conserved in the operator sequences of other bacterial regulators such as GntR (6) and PdhR (27) (which belong to the GntR family) and, surprisingly, CRP (23, 25) and LacI (26) (which belong to the GalR-LacI family), the RH region differs from that of other operators (Fig. 8). Interestingly, the TGTGA subsequence of the LH region in $OR_A$ is also present in the LH region of the LacI- and CRP-binding sites (Fig. 8), where it has been shown to play a critical role in binding to the cognate regulators (25, 28). In this sense, we have shown here that mutations in the LH region of $OR_A$ are more deleterious for PaaX binding and repression *in vivo* than mutations in the RH region, suggesting that the repressor makes rather different sets of contacts with the two halves of the binding site and the key interactions are those with the LH region. Similar non-identical half-site interactions have been observed also within the *lac* operator; again the LH region (the one containing the TGTGA subsequence) is the most critical for the interactions with the LacI repressor (28). The observed DNase I-hypersensitive sites in the RH region of $OR_A$ and $OR_Z$ could also reflect a more complete protection of the LH regions by the PaaX repressor, a situation that has been already reported for other operator sequences (28). Although the conserved PaaX binding motif is essential for PaaX binding, point mutations outside this motif, such as that of the mutant promoter *PaM22*, seem to influence the intrinsic affinity of the repressor for the promoter. Similar data showing that regions distal to the regulator binding site contribute slightly to the interactions with the cognate proteins have been previously reported for other regulators (25, 26, 29).

Since induction of the gene expression driven by the *Pa* (2) and *Pz* (Table II) promoters requires both PA and a functional PA-CoA ligase, it was reasonable to anticipate that the true inducer of the *paa* catabolic cluster was the first intermediate of the pathway, *i.e.* PA-CoA, rather than the initial substrate, *i.e.* PA. This was confirmed *in vitro* through gel retardation assays which revealed that binding of PaaX to the *Pa* and *Pz* promoters is specifically prevented by PA-CoA and not by PA (Fig. 2), benzoyl-CoA, acetyl-CoA, or free CoA. To date, the *E. coli* FadR protein, which controls the expression of many genes involved in fatty acid synthesis and degradation, was the only regulator for which there is convincing evidence that interaction of alkyl-CoA compounds with the protein prevents DNA binding (30). The PaaX repressor constitutes, therefore, the first example of a regulator protein that responds specifically to an aryl-CoA compound. It is likely that other regulatory proteins showing a significant amino acid sequence similarity with PaaX, *e.g.* the equivalent PaaN repressor (41.4% identity) of the PA degradation pathway from *P. putida* U (1), the putative regulator, ORF13 (31.5% identity), of a predicted PA catabolic pathway from *Bacillus halodurans* (GenBank™ accession number AB011837), and the putative repressor (GenBank™ accession number AF042490) (24.1% identity) of the 4-chlorobenzoate dehalogenation pathway involving formation of CoA derivatives in *Arthrobacter* sp. TM1 (31), also interact with aryl-CoA compounds.

In addition to the operon-specific regulation mediated by the PaaX repressor, the *paa* catabolic cluster of *E. coli* is subjected to control by the two global regulators, CRP (23) and IHF (22). We have observed that $CRP^-$ *E. coli* strains failed to express the *Pa-lacZ* and *Pz-lacZ* translational fusions, indicating that CRP acts as an activator of the gene expression driven by the *Pa* and *Pz* promoters. Gel retardation assays confirmed the binding of the cAMP·CRP complex to the *Pa* promoter region (Fig. 6). Since a potential CRP-binding site was identified at position −61.5 with respect to the major transcription start site of *Pa* (Fig. 4), this promoter might follow a CRP-dependent activation mechanism similar to that described for class I promoters (32). Although CRP is also necessary for activity of the *Pz* promoter, no binding of CRP to the Pz fragment was observed in gel retardation assays (Fig. 6), suggesting that CRP bound to *Pa* is able to activate the divergent *Pz* promoter. Since the putative CRP-binding site is at position −138.5 with respect to the transcription start site of *Pz* and, therefore, too far upstream to activate directly transcription from this promoter (33), additional transcription factors may be involved. We have observed that IHF binds to the *paaZ-paaA* intergenic region and stimulates transcription from the *Pa* and *Pz* promoters (Fig. 7). Nucleotide sequence analysis revealed a putative IHF-binding site within the Pz fragment at position −115 with respect to the major transcription start site of *Pa* (Fig. 4). Whether IHF-induced bending of the *paaZ-paaA* intergenic region can bring together RNA polymerase bound to *Pz* and CRP bound to *Pa* could be an interesting model that remains to be tested.

Carbon catabolite repression has been extensively described for the degradation of aromatic compounds in different bacteria (34). Here we report that when *E. coli* cells are grown in

PA-containing minimal medium in the presence of their preferred carbon source, *i.e.* glucose, gene expression from the *Pa* and *Pz* promoters is subject to carbon catabolite repression. In contrast to the situation in *Pseudomonas* and other Gram-negative bacteria where the mechanism of catabolite repression is not yet understood (34, 35), the effect of glucose in *E. coli* is mediated through the action of CRP (23, 25). A similar repression effect of glucose mediated by the cAMP·CRP complex was also observed in *E. coli* for the catabolism of other aromatic compounds such as 2-phenylethylamine (36) and 4-hydroxyphenylacetic acid (14), and it is likely to occur also for homoprotocatechuic acid degradation (7, 8). Interestingly, the catabolism of PA in *P. putida* U is also under catabolic repression by glucose (1). It is well known that promoters are subjected to various types of physiological controls which adjust their transcriptional output to the general environmental conditions of the cells (34). We have shown in this work that the specific PaaX-mediated control of the *Pa* and *Pz* promoters for PA degradation in *E. coli* is, in turn, subordinated to a superimposed regulation mediated by global regulators such as CRP and IHF, which connect the expression of the *paa* catabolic genes to the metabolic and energetic status of the cell.

*Acknowledgments*—We thank M. K. B. Berlyn for strain SH210; V. de Lorenzo for strains S90C, DPB101, and DBP102; and J. M. Luengo for the CoA derivatives. We are indebted to S. Goodman and A. Kolb for the kind gift of purified IHF and CRP, respectively. We gratefully acknowledge the help of E. Aporta with oligonucleotide synthesis; A. Díaz, G. Porras, and S. Carbajo with sequencing; and the technical assistance of E. Cano, M. Carrasco, and F. Morante. We thank G. Bertoni, L. Escolar, and C. Pérez for their advice with the gel retardation and footprinting assays. We are indebted to V. de Lorenzo for the inspiring discussions and to M. A. Prieto and C. Harwood for the critical reading of the manuscript.

REFERENCES

1. Olivera, E. R., Miñambres, B., García, B., Muñiz, C., Moreno, M. A., Ferrández, A., Díaz, E., García, J. L., and Luengo, J. M. (1998) *Proc. Natl. Acad. Sci. U. S. A.* **95,** 6419–6424
2. Ferrández, A., Miñambres, B., García, B., Olivera, E. R., Luengo, J. M., García, J. L., and Díaz, E. (1998) *J. Biol. Chem.* **273,** 25974–25986
3. Ferrández, A., Prieto, M. A., García, J. L., and Díaz, E. (1997) *FEBS Lett.* **406,** 23–27
4. García, B., Olivera, E. R., Miñambres, B., Fernández-Valverde, M., Cañedo, L. M., Prieto, M. A., García, J. L., Martínez, M., and Luengo, J. M. (1999) *J. Biol. Chem.* **274,** 29228–29241
5. Raman, N., Black, P. N., and DiRusso, C. C. (1997) *J. Biol. Chem.* **272,** 30645–30650
6. Fujita, Y., and Miwa, Y. (1989) *J. Biol. Chem.* **264,** 4201–4206
7. Prieto, M. A., Díaz, E., and García, J. L. (1996) *J. Bacteriol.* **178,** 111–120
8. Roper, D. I., Fawcett, T., and Cooper, R. A. (1993) *Mol. Gen. Genet.* **237,** 241–250
9. Eaton, R. W. (1997) *J. Bacteriol.* **179,** 3171–3180
10. Arai, H., Akahira, S., Ohishi, T., and Kudo, T. (1999) *Mol. Microbiol.* **33,** 1132–1140
11. Sambrook, J., Fritsch, E. F., and Maniatis, T. (1989) *Molecular Cloning: A Laboratory Manual,* 2nd Ed., Cold Spring Harbor Laboratory, Cold Spring Harbor, NY
12. de Lorenzo, V., and Timmis, K. N. (1994) *Methods Enzymol.* **235,** 386–405
13. Biek, D., and Cohen, S. N. (1989) *J. Bacteriol.* **171,** 2056–2065
14. Prieto, M. A., and García, J. L. (1997) *Biochem. Biophys. Res. Commun.* **232,** 759–765
15. Schweizer, H. P., and Boos, W. (1983) *Mol. Gen. Genet.* **192,** 293–294
16. Miller, J. H. (1972) *Experiments in Molecular Genetic*s, Cold Spring Harbor Laboratory, Cold Spring Harbor, NY
17. Sanger, F., Nicklen, S., and Coulson, A. R. (1977) *Proc. Natl. Acad. Sci. U. S. A.* **74,** 5463–5467
18. Bradford, M. M. (1976) *Anal. Biochem.* **72,** 248–254
19. Maxam, A, and Gilbert, W. (1980) *Methods Enzymol.* **65,** 499–560
20. Azizan, A., Sherin, D., DiRusso, C. C., and Black, P. N. (1999) *Arch. Biochem. Biophys.* **365,** 299–306
21. Pabo, C. O., and Sauer, R. T. (1992) *Annu. Rev. Biochem.* **61,** 1053–1095
22. Goosen, N., and van de Putte, P. (1995) *Mol. Microbiol.* **16,** 1–7
23. Saier, M. H., Ramseier, T. M., and Reizer, J. (1996) in *Escherichia coli and Salmonella: Cellular and Molecular Biology* (Neidhardt, F. C., Curtiss, R., III, Ingraham, J. L., Lin, E. C. C., Low, K. B., Magasanik, B., Reznikoff, W. S., Riley, M., Schaechter, M., and Umbarger, H. E., eds) pp. 1325–1343, American Society for Microbiology, Washington, D. C.
24. Vicente, M., Chater, K. F., and de Lorenzo, V. (1999) *Mol. Microbiol.* **33,** 8–17
25. Kolb, A., Busby, S., Buc, H., Garges, S., and Adhya, S. (1993) *Annu. Rev. Biochem.* **62,** 749–795
26. Rastinejad, F., Artz, P., and Lu, P. (1993) *J. Mol. Biol.* **233,** 389–399
27. Quail, M. A., and Guest, J. R. (1995) *Mol. Microbiol.* **15,** 519–529
28. Sadler, J. R., Sasmor, H., and Betz, J. L. (1983) *Proc. Natl. Acad. Sci. U. S. A.* **80,** 6785–6789
29. Yoshida, K.-I., Shibayama, T., Aoyama, D., and Fujita, Y. (1999) *J. Mol. Biol.* **285,** 917–929
30. DiRusso, C. C., Tsvetnitsky, V., Hojrup, P., and Knudsen, J. (1998) *J. Biol. Chem.* **273,** 33652–33659
31. Schmitz, A., Gartemann, K.-H., Fiedler, J., Grund, E., and Eichenlaub, R. (1992) *Appl. Environ. Microbiol.* **58,** 4068–4071
32. Busby, S., and Ebright, R. H. (1999) *J. Mol. Biol.* **293,** 199–213
33. Ushida, C., and Aiba, H. (1990) *Nucleic Acids Res.* **18,** 6325–6330
34. Cases, I., and de Lorenzo, V. (1998) *Curr. Opin. Microbiol.* **1,** 303–310
35. MacGregor, C. H., Wolff, J. A., Arora, S. K., Hylemon, P. B., and Phibbs, P. V. (1992) in *Pseudomonas Molecular Biology and Biotechnology* (Galli, E., Silver, S., and Witholt, B., eds) pp. 198–206, American Society for Microbiology, Washington, D. C.
36. Yamashita, M., Azakami, H., Yokoro, N., Roh, J.-H., Suzuki, H., Kumagai, H., and Murooka, Y. (1996) *J. Bacteriol.* **178,** 2941–2947

# Transcriptional Regulation of the Divergent *paa*Catabolic Operons for Phenylacetic Acid Degradation in *Escherichia coli*

Abel Ferrández, José Luis Garci?a and Eduardo Di?az

*J. Biol. Chem. 2000, 275:12214-12222.*
*doi: 10.1074/jbc.275.16.12214*

---